

# Strong constraints on self-interacting dark matter with light mediators

Torsten Bringmann\*

*Department of Physics, University of Oslo, Box 1048, N-0371 Oslo, Norway*

Felix Kahlhoefer<sup>†</sup> and Kai Schmidt-Hoberg<sup>‡</sup>

*Deutsches Elektronen-Synchrotron DESY, Notkestraße 85, D-22607 Hamburg, Germany*

Parampreet Walia<sup>§</sup>

*Department of Physics, University of Oslo, Box 1048, N-0371 Oslo, Norway*

Coupling dark matter to light new particles is an attractive way to combine thermal production with strong velocity-dependent self-interactions. Here we point out that in such models the dark matter annihilation rate is generically enhanced by the Sommerfeld effect, and we derive the resulting constraints from the Cosmic Microwave Background and other indirect detection probes. For the frequently studied case of s-wave annihilation these constraints exclude the entire parameter space where the self-interactions are large enough to address the small-scale problems of structure formation.

*Introduction.*— Although dark matter (DM) particles can only have very weak interactions with Standard Model (SM) states, it is an intriguing possibility that they experience much stronger self-interactions. Such self-interactions may alter the behaviour of DM on astrophysical and cosmological scales in striking ways. In particular, self-interacting DM (SIDM) may offer an attractive solution to some of the long-standing small-scale structure problems encountered in the collisionless cold DM paradigm [1, 2]. At the same time, a conclusive observation of collisionality of DM on astrophysical scales would have striking implications for the particle physics properties of DM. For these reasons, SIDM has been the subject of increasing interest in the last few years.

In order to affect astrophysical observations, the DM self-scattering cross section typically has to be of order  $\sigma/m_\chi \sim 1 \text{ cm}^2 \text{ g}^{-1}$  [3–6]. Such large cross sections can arise from the underlying particle physics in essentially two different ways: either from new strong forces in the dark sector, similar to QCD, or from a more weakly coupled theory with a very light mediating particle. In the former case, large self-interactions are expected for all DM velocities, leading to strong bounds in particular from galaxy clusters [7–12]. In contrast, in models with a very light mediator self-interactions become stronger at smaller DM velocities, so that large effects on small scales can be consistent with the stronger astrophysical constraints on larger scales [3, 4, 13–16].

Another attractive feature of SIDM with a very light mediator is that the DM abundance today can be explained by thermal production in the early Universe. In contrast to the standard WIMP scenario [17], where DM annihilates directly into SM states (see [18]), the DM particle  $\chi$  annihilates into the mediator particles  $\phi$ , which subsequently decay into SM states:  $\chi\chi \rightarrow \phi\phi \rightarrow \text{SM}$  [19]. This can naturally yield the observed relic abundance even if interactions with the SM are strongly suppressed.

From the particle physics perspective, the presence of

very weakly coupled light mediators can be easily motivated, for example if the stability of the DM particle arises from a charge under a new spontaneously broken  $U(1)'$  gauge group. Nevertheless, the mediator should couple to the SM at *some* level [20, 21] in order to *i*) achieve thermal equilibrium between the dark sector and the SM in the early Universe, and *ii*) to decay sufficiently quickly to avoid overclosing the Universe. Such scenarios provide a both attractive and minimal realization of the SIDM idea, and have hence been one of the main avenues of model building [22–24].

The required interaction strength between mediators and the SM turns out to be rather small, and is therefore usually assumed to be essentially unconstrained. In this Letter we demonstrate that decays of the mediator into SM states are, instead, very strongly constrained for such models. The reason is that the large self-interaction cross sections are achieved via non-perturbative enhancements which at the same time also enhance the DM annihilation cross section, in particular for small DM velocities [25–28]. The subsequent decays of the mediators into SM states would then typically change the reionization history of the Universe, and thereby lead to significant distortions of the Cosmic Microwave Background (CMB), or be observed in indirect detection experiments today.

This Letter is organized as follows. We first revisit the properties of the DM and mediator particles that are required to reproduce the observed relic abundance and give rise to phenomenologically relevant DM self-interactions. We then discuss CMB and indirect detection constraints, with special emphasis on strongly velocity-dependent DM annihilation rates. We illustrate their impact on the most popular class of models, in which the DM relic density is set by s-wave annihilation. Finally, we comment on how the resulting strong constraints may be relaxed.

*Self-interacting DM with light mediators.*— We consider a non-relativistic DM species  $\chi$  that interacts via

a light vector or scalar mediator  $\phi$ . The DM self-interactions that result from exchanging  $\phi$  can be described by a Yukawa potential, which leads to a strong dependence of the self-interaction rate on the relative velocity  $v$  of the scattering DM particles. The phenomenology of this scenario is fully characterized by the two masses,  $m_\chi$  and  $m_\phi$ , and the coupling strength  $\alpha_\chi \equiv g^2/4\pi$ .

In the *Born limit* ( $\alpha_\chi m_\chi \lesssim m_\phi$ ), the momentum transfer cross section  $\sigma_T$  can be calculated perturbatively [29]. For larger coupling strengths or DM masses, non-perturbative effects become important. In the following, we use the improved parameterization from [30] for the *classical limit* ( $m_\chi v \gtrsim m_\phi$ ) and adopt the analytical expressions from [16], which have been obtained from approximating the Yukawa potential by a Hulthén potential, in the *intermediate (resonant) regime*. To estimate the effect of DM self-interactions on dwarf galaxies, we define  $\langle\sigma_T\rangle_{30}$  as  $\sigma_T$  averaged over a Maxwellian velocity distribution with a most probable velocity of  $30 \text{ km s}^{-1}$ . To obtain observationally relevant effects, e.g. to alleviate the cusp-core [31–33] and too-big-to-fail [34, 35] problems, we require  $\langle\sigma_T\rangle_{30}/m_\chi \sim 0.1\text{--}10 \text{ cm}^2 \text{ g}^{-1}$  [5, 6].

As motivated in the introduction, a second important constraint can be obtained under the assumption that the dark sector was in thermal equilibrium with the SM sector at early times and the DM relic abundance is set by thermal freeze-out. This assumption is well-motivated if the mediator couples also to SM states, but may need to be revisited if these interactions are very weak (see below). As the dominant DM annihilation channel is  $\chi\chi \rightarrow \phi\phi$ , we can effectively eliminate  $\alpha$  as a free parameter by requiring that the relic density matches the observed value of  $\Omega_\chi h^2 = 0.1188 \pm 0.0010$  [36].

The required value of  $\alpha$  depends on the particle masses and on whether the annihilation proceeds via an *s*- or a *p*-wave process. The former implies a constant annihilation rate  $(\sigma v)_0$  at the perturbative level, for  $v \ll 1$ , while the latter implies  $(\sigma v)_0 \propto v^2$ . In both cases, we can combine the requirement of sizeable self-interaction rates with the observed relic density for the three regimes mentioned above [16]. As visualized later in Fig. 2, this yields very roughly ( $m_\chi \gtrsim 100 \text{ GeV}$ ,  $m_\phi \lesssim 10 \text{ MeV}$ ) in the classical regime, ( $10 \text{ GeV} \lesssim m_\chi \lesssim 100 \text{ GeV}$ ,  $1 \text{ MeV} \lesssim m_\phi \lesssim 1 \text{ GeV}$ ) in the resonant regime, and ( $m_\chi \lesssim 10 \text{ GeV}$ ,  $m_\phi \lesssim 10 \text{ MeV}$ ) in the Born limit.

*Sommerfeld enhancement.*— The Yukawa potential due to light mediator exchange does not only affect DM self-interactions, but it also modifies the wave-function of the annihilating DM pair [37, 38]. For small velocities, this can lead to significant non-perturbative corrections to the tree-level annihilation rate,  $\sigma v = S \times (\sigma v)_0$ , with the Sommerfeld enhancement factor  $S$  given in [39–41]. For  $\alpha_\chi m_\phi \ll m_\chi v^2$ , the Yukawa potential becomes indistinguishable from a Coulomb potential and no strong resonances appear in  $S$ .

This effect is usually taken into account for relic density calculations in SIDM models, and we adopt here the results from [15, 16]. Quantitatively, the required value

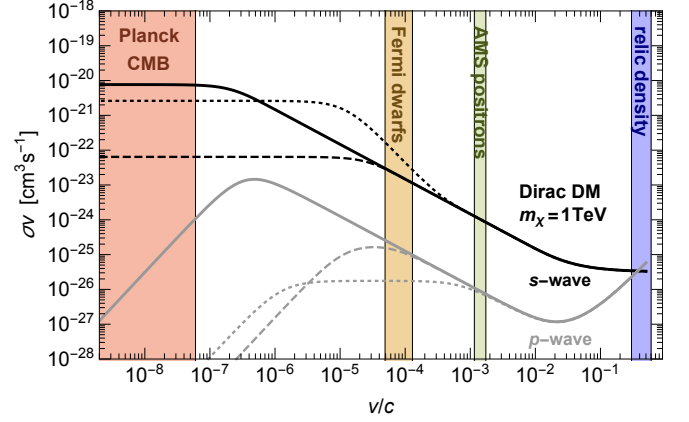


FIG. 1. Comparison of cross sections for *s*-wave and *p*-wave annihilation, as a function of the relative DM-DM velocity. The coupling  $\alpha_\chi$  is fixed by the relic density requirement. Solid (dashed) curves correspond to  $m_\chi = 1 \text{ TeV}$  and  $m_\phi = 1 \text{ MeV}$  ( $m_\phi = 100 \text{ MeV}$ ), while dotted lines show the case of  $m_\phi$  tuned to  $1.112 \text{ GeV}$  ( $1.066 \text{ GeV}$ ) for resonant *s*-wave (*p*-wave) annihilation. In addition, the typical velocity ranges of different experimental probes are indicated.

of  $\alpha_\chi$  differs from the perturbative result only by an  $\mathcal{O}(1)$  factor independent of  $m_\phi$ , because for most of the parameter space of interest we are in the Coulomb regime during chemical freeze-out. We neglect the model-dependent effect of a second period of DM annihilation *after kinetic decoupling* that can occur for Sommerfeld-enhanced DM annihilation [26, 27, 42, 43]. While this may in principle decrease  $\Omega_\chi$  by up to three orders of magnitude if DM annihilation occurs very close to a resonance, it changes the calculation only at the percent level off resonance [43]. Similarly, we neglect the effect of bound-state formation, which only becomes important close to the unitarity bound [44].

As the Universe continues to cool down after DM freeze-out, the DM velocities decrease. The crucial observation for the purpose of this Letter is that the Sommerfeld enhancement at late times is therefore much larger than during freeze-out. We illustrate this in Fig. 1, where we show *s*-wave and *p*-wave annihilation cross sections as a function of the DM velocity for different mediator masses, with  $\alpha_\chi$  being fixed by the relic density requirement. Away from any resonance, the enhancement scales like  $1/v$  and  $1/v^3$  for the *s*-wave and *p*-wave case, respectively, so that effectively the cross sections scale like  $1/v$  in both cases. For the *p*-wave case, however, there is an offset compared to the thermal cross section due to the initial  $v^2$  suppression. In both cases the saturation of the Sommerfeld enhancement occurs at about  $v \sim m_\phi/2m_\chi$ , leading to a plateau for the *s*-wave and a maximum for the *p*-wave case. For masses tuned to a resonance (as shown for the dotted lines), the *s*-wave enhancement grows like  $1/v^2$  and saturates later.

This figure clearly demonstrates that probes of DM annihilation at small velocities have the potential to seri-

ously impact SIDM models with light mediators, in particular for  $s$ -wave annihilation. In the following, we will discuss the relevant observational constraints in turn and then quantify our general expectation by considering a popular concrete model realization that leads to  $s$ -wave dominated DM annihilation.

*CMB constraints.*— Mediator particles decaying into SM particles at a rate larger than the Hubble rate can bring the model in conflict with the robust constraints on DM annihilation from CMB observations [45–50]. At 95% C.L. the most recent Planck data result in [36]

$$\frac{\langle\sigma v\rangle_{\text{rec}}}{N_\chi} \lesssim 4 \times 10^{-25} \text{ cm}^3 \text{ s}^{-1} \left( \frac{f_{\text{eff}}}{0.1} \right)^{-1} \left( \frac{m_\chi}{100 \text{ GeV}} \right), \quad (1)$$

where  $\langle\sigma v\rangle_{\text{rec}}$  is the annihilation rate at recombination ( $z_{\text{rec}} \sim 1100$ ) averaged over the distribution of DM velocities and  $N_\chi = 1$  ( $N_\chi = 2$ ) for Majorana (Dirac) DM. The efficiency factor  $f_{\text{eff}}$  is related to the fraction of the released energy ending up in photons or electrons (see e.g. [51]) with  $f_{\text{eff}} \gtrsim 0.1$  for any SM final state apart from neutrinos.

The typical DM velocity  $v_\chi$  during recombination depends on the temperature of kinetic decoupling  $T_{\text{kd}}$  via  $\langle v_\chi^2 \rangle_{\text{rec}} = \frac{1}{2} \langle v^2 \rangle_{\text{rec}} = 3 (T_{\text{kd}}/m_\chi) (z_{\text{rec}}/z_{\text{kd}})^2$  [52], where  $z_{\text{kd}}$  is the redshift at kinetic decoupling. For standard WIMPs, one expects  $10 \text{ MeV} \lesssim T_{\text{kd}} \lesssim 1 \text{ GeV}$  [52]. In the presence of light mediators, however, kinetic decoupling can be significantly delayed [43]. Nevertheless, observations of the Lyman- $\alpha$  forest [53, 54] robustly exclude  $T_{\text{kd}} \lesssim 100 \text{ eV}$  (see Ref. [55] for a recent discussion). This translates into an upper bound on the relative DM velocity at recombination of

$$v_{\text{rec}} \lesssim 2 \times 10^{-7} \left( \frac{m_\chi}{100 \text{ GeV}} \right)^{-1/2}. \quad (2)$$

Such small velocities imply enormously enhanced DM annihilation rates.

As an illustrative example, let us estimate the effect on the classical regime of DM self-scattering, assuming that annihilation proceeds via an  $s$ -wave process. Since there are no resonances, the Sommerfeld enhancement saturates for  $v \lesssim m_\phi/2m_\chi$ . For phenomenologically relevant values of  $\langle\sigma_T\rangle_{30}$ , this mass ratio is numerically larger than the minimal velocity in Eq. (2). The annihilation rate relevant for CMB constraints thus becomes *maximal*,

$$\langle\sigma v\rangle_{\text{rec}} \sim \langle\sigma v\rangle_{\text{cd}} v^* \frac{m_\chi}{m_\phi}, \quad (3)$$

where the annihilation rate at freeze-out is approximately  $\langle\sigma v\rangle_{\text{cd}}/N_\chi \sim 3 \times 10^{-26} \text{ cm}^3 \text{ s}^{-1}$  in order to obtain the correct relic density and  $v^* \sim 0.1$  denotes the velocity below which  $\sigma v \propto 1/v$ . Comparing Eq. (1) to Eq. (3), we conclude that for  $s$ -wave annihilation in the classical scattering regime this class of models is ruled out unless either  $m_\phi \gtrsim 1 \text{ GeV}$ , in which case no sizeable DM self-interactions can be achieved, or  $f_{\text{eff}} \ll 0.1$ .

*Other indirect detection constraints.*— In the Born and resonant regimes there are additional strong constraints from observations probing somewhat larger DM velocities, specifically from searches for present-day DM annihilation into light mediators  $\phi$ , which in turn decay into SM particles. For simplicity, we will here only consider light leptonic decay modes, noting that in general further channels will open up for  $m_\phi \gtrsim 2m_{\pi^0}$  and the resulting constraints would be more stringent.

Dwarf galaxy observations with the Fermi gamma-ray space telescope provide one of the most robust ways to constrain DM annihilation, and we implement them using the likelihood functions provided by the Fermi-LAT collaboration which extend down to photon energies of 500 MeV [56]. To obtain the gamma-ray spectrum, we first calculate the distribution of photon energies from  $\phi \rightarrow \ell^+ \ell^- \gamma$ , in the rest-frame of  $\phi$ , and then boost it to the DM frame as in [57]. Kinematical data constrain the DM velocities to be much smaller than the  $v_\chi \sim 10^{-3}$  observed in our Galaxy; here, we adopt a relative velocity of  $v = 10^{-4}$ , which independently of the assumed profile is a conservative choice [58]. We use the  $J$ -factors assuming a Burkert profile from [59].

Local DM annihilation to positrons are strongly constrained by the high-accuracy data of the AMS-02 experiment [60, 61], with only moderate uncertainties related to the local normalizations of the DM profile and radiation density [62]. We take the bounds from [63] for one-step cascade annihilations with the intermediate state decaying to  $e^+ e^-$  and  $\mu^+ \mu^-$ . These bounds extend down to DM masses of 10 GeV and, for  $m_\phi \ll m_\chi$ , are to good approximation independent of the mediator mass.

*Example model.*— Let us now apply the above constraints to an often discussed example of a model with  $s$ -wave DM annihilations [14, 16, 20, 38, 64]. Here,  $\chi$  is a Dirac fermion that couples to a massive vector  $\phi$ . The latter can obtain couplings to SM particles via kinetic mixing with the hypercharge field strength  $B^{\mu\nu}$  or via mass mixing with the  $Z$  [18, 65, 66]:

$$\mathcal{L} \supset -g_\chi^\nu \phi^\mu \bar{\chi} \gamma_\mu \chi - \frac{1}{2} \sin \epsilon B_{\mu\nu} \phi^{\mu\nu} - \delta m^2 \phi^\mu Z_\mu. \quad (4)$$

We first focus on the case of negligible mass mixing,  $\delta m \ll m_\phi$ . For  $m_\phi \ll m_Z$  the couplings of the mediator are then largely photon-like, so this situation is very similar to kinetic mixing with electromagnetism [67, 68]. The dominant decay mode for  $m_\phi \lesssim 1 \text{ MeV}$  is then  $\phi \rightarrow 3\gamma$ , which is so small that  $\phi$  would effectively be stable; in such a scenario the mediator would either overclose the Universe or, close to the threshold, produce gamma rays in excess of the extragalactic gamma-ray background [68]. For heavier mediator masses, decays into leptons and hadrons open up, for which we adopt the branching ratios from [69].

For each combination of DM and mediator mass in this model, we calculate the Sommerfeld enhancement factor using the conservative upper bound on  $v_{\text{rec}}$  from Eq. (2). By comparing the result to Eq. (1), we can determine

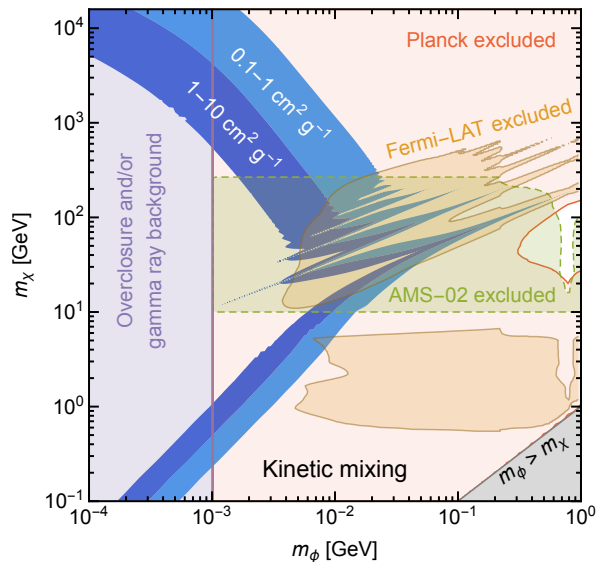


FIG. 2. Constraints at 95% C.L. on DM annihilating into vector mediators that kinematically mix with hypercharge as a function of the DM and mediator masses. The blue shaded region shows the combinations of DM mass  $m_\chi$  and mediator mass  $m_\phi$  that lead to a DM self-interaction cross section of  $0.1 \text{ cm}^2 \text{ g}^{-1} < \langle \sigma_T \rangle_{30} / m_\chi < 10 \text{ cm}^2 \text{ g}^{-1}$ , which would visibly affect astrophysical observables at dwarf galaxy scale [16].

the parameter region excluded by CMB constraints. To calculate the appropriate value of  $f_{\text{eff}}$  as a function of  $m_\chi$  and  $m_\phi$ , we multiply the different decay modes with the efficiency factors from [51]. Our results are shown in Fig. 2, where we also show the Fermi and AMS-02 bounds discussed above. We observe that the CMB constraints, and partially also the other indirect detection constraints, exclude all combinations of  $m_\chi$  and  $m_\phi$  that lead to interesting self-interaction cross sections.

We emphasize that very close to a resonance both the preferred SIDM region and the various constraints may be modified by the impact of a potential second period of DM annihilation on the relic density calculation (see above). For late kinetic decoupling the resulting modifications will be small, but we expect even larger effects not to change our results qualitatively.

*Discussion.*— The bounds shown in Fig. 2 have been obtained under very conservative assumptions and are expected to apply in a similar way to other models of spin-1 mediators. The CMB constraints, in particular, are very robust because we probe DM annihilation in a kinematical situation where the Sommerfeld enhancement is typically already saturated. Even for parameter combinations where this is not the case, our constraints are extremely conservative because we evaluate  $\sigma v$  no later than at recombination, and for larger values of  $v_{\text{rec}}$  than expected in a realistic treatment of kinetic decoupling. Nevertheless, our analysis does rely on a number of assumptions, which we will now review in detail.

For our calculations so far, there was no need to specify

the kinetic mixing parameter  $\epsilon$ , as long as mixing is sufficiently large that the mediator decays in time to affect the reionisation history. Nevertheless, we have assumed implicitly that  $\epsilon$  is large enough to thermalise the visible sector and the dark sector before freeze-out. Depending on the DM mass, the required value of  $\epsilon$  for this to happen is of order  $10^{-7}$ – $10^{-5}$  [70]. However, DM direct detection experiments (as well as astrophysical constraints for  $m_\phi \lesssim 1 \text{ MeV}$  [71]) typically require much smaller values of  $\epsilon$  [20]. The conclusion is that a different mechanism must be responsible for bringing the visible and the dark sector into thermal contact.

The simplest possibility would be a thermal contact at higher temperatures, via a different portal. After this interaction ceases to be effective, the temperatures of both sectors would then evolve independently, depending on the number of degrees of freedom in each sector. For sizeable  $\alpha_\chi$  the DM relic abundance will still be determined by dark sector freeze out, but at a different temperature. For reasonable temperature ratios, as we discuss in detail in Appendix A, such a situation does not lead to qualitatively different results compared to the case where the two sectors have the same temperature. For the case where the two sectors never reach thermal equilibrium and the DM relic abundance is for example set via the freeze-in mechanism, we refer to [23].

A second important assumption is that DM can self-annihilate via an  $s$ -wave process. In the  $p$ -wave case the resulting cross sections are much smaller, even though the Sommerfeld enhancement can be quite significant (see Fig. 1). Moreover, at very small velocities the cross section again decreases like  $v^2$  and therefore typically becomes unobservably small at recombination. Not surprisingly, we therefore find that both CMB and indirect detection bounds can be evaded for most of the parameter space. However, models leading to  $p$ -wave annihilation are strongly constrained from independent model building considerations, in particular the combination of constraints from direct detection experiments and primordial nucleosynthesis [20].

Finally, our conclusions can be modified if the mediator decays in a different way than via kinetic mixing. As a specific example, we discuss the case of mass mixing in Appendix B, in which case the mediator obtains a significant coupling to neutrinos. This alleviates constraints from both DM annihilation and the mediator lifetime, but in principle offers exciting prospects for indirect detection [15]: DM annihilation into a pair of mediators followed by the decay  $\phi \rightarrow \bar{\nu}\nu$  would result in a characteristic spectral feature [72]. While currently unconstrained for the models considered here, such a signal is in reach for IceCube observations of the Galactic halo [73–76].

In general, however, the constraints derived above are so strong that they can even be applied to models where mediator decays into leptons are sub-dominant. As a result, large self-interactions are excluded also for the case of mass mixing, as long as  $m_\phi > 2m_e$ . Even weaker constraints could in principle be obtained if the media-

tors couple to another very light state in the dark sector, such as sterile neutrinos. Such models are particularly interesting because they can significantly delay kinetic decoupling and thus provide a solution also to the missing satellite problem [15, 55, 77–81].

*Conclusions.*— Models of DM with velocity-dependent self-interactions have recently received a great deal of attention for their potential to produce a number of interesting effects on astrophysical scales. We have shown in this Letter that these models face very strong constraints from the CMB and DM indirect detection. In the most natural realization of this scenario with a light vector mediator with kinetic mixing, these constraints rule out the entire parameter space where the self-scattering cross section can be relevant for astrophysical systems. These bounds remain highly relevant for a number of generalizations of the scenario, such as a different dark sector temperature and different mediator branching ratios. Clearly, future efforts to develop particle physics models for SIDM need to address these issues in order to arrive at models that provide a picture consistent with all observations in cosmology, astrophysics and particle physics.

*Acknowledgements.*— We thank Camilo Garcia-Cely, Michael Gustafsson, Julian Heeck, Andrzej Hryczuk, Joerg Jaeckel, Andreas Ringwald, Marco Taoso, Sebastian Wild and Bryan Zaldivar for enlightening discussions. This work is supported by the German Science Foundation (DFG) under the Collaborative Research Center (SFB) 676 Particles, Strings and the Early Universe as well as the ERC Starting Grant ‘NewAve’ (638528).

## Appendix A: Modifying the dark sector temperature

If the couplings of the new mediator to the SM sector become very small, the dark and visible sectors will no longer be able to thermalize with each other in the early Universe. In principle, the temperature of the dark sector can then be almost arbitrary and could for example be set by the details of reheating. Nevertheless, it makes sense to assume that some other mechanism (e.g. another *heavy* mediator) brings the two sectors into thermal equilibrium at high temperatures. Once this mechanism becomes inefficient, the temperatures of the two sectors will evolve independently and will in general differ due to the different number of degrees of freedom in the two sectors.

Denoting the ratio of effective relativistic degrees of freedom in the dark sector and the visible sector by  $\eta \equiv g_{*,\text{dark}}/g_{*,\text{vis}}$ , the temperature ratio between the two sectors,  $\xi \equiv T_{\text{dark}}/T_{\text{vis}}$ , is given by  $\xi(T) = [\eta(T_{\text{dec}})/\eta(T)]^{1/3}$  as long as entropy is conserved separately in the two sectors. Here,  $T$  denotes the standard photon temperature, and  $T_{\text{dec}}$  its value right after decoupling. Since the number of effective degrees of freedom in the visible sector typically decreases more rapidly with decreasing temperature, in particular if one

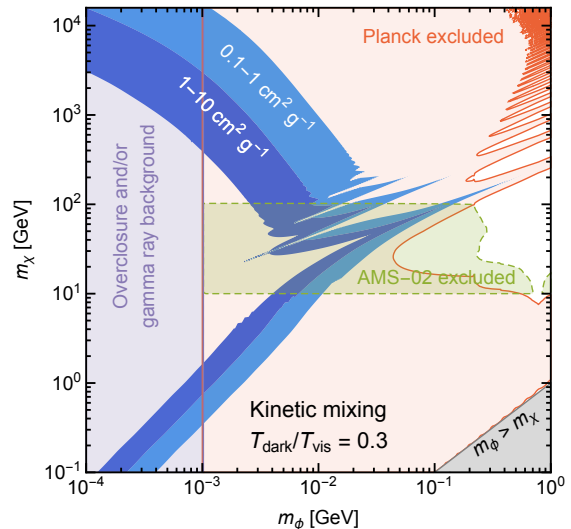


FIG. 3. Constraints on DM particles annihilating into vector mediators that kinematically mix with hypercharge, assuming a temperature ratio of  $\xi = 0.3$  between dark and visible sector during the thermal freeze-out of DM.

does not want to proliferate the number of new states, this places a lower bound on the temperature ratio. If the mediator and DM particles are the only new states in the dark sector right after decoupling, for example,  $g_{*,\text{dark}}$  does not change at all and this lower bound is given by  $\xi(T) \gtrsim 0.34 [g_{*,\text{vis}}(T_{\text{dec}})/100]^{-1/3}$  when conservatively assuming that the thermal freeze-out of DM happens as late as at  $T_{\text{cd}} \lesssim 1$  MeV [55] (in reality, we have  $T_{\text{cd}} \sim m_\chi/25$ , leading to considerably larger lower bounds on  $\xi$  for DM heavier than about 100 MeV).

This change in temperature changes the value of  $\alpha_\chi$  that gives the correct relic abundance. To first approximation, and neglecting additional mass-dependent corrections, one can replace  $\alpha_\chi \rightarrow \sqrt{\xi}\alpha_\chi$  [55, 82]. Since the late time annihilation and the self scattering cross sections scale differently with  $\alpha_\chi$ , the constraints can be somewhat alleviated, as shown in Fig. 3 for the rather extreme case of  $\xi = 0.3$ .

## Appendix B: Mass mixing

Let us now consider the case that the decay modes of the mediator are dominantly set by its mass mixing with the SM  $Z$  boson, which requires  $\delta m/m_Z \gg \epsilon$ . The most important change in this case is that the mediator obtains sizeable couplings to neutrinos. As a result, the overclosure and gamma-ray constraints for  $m_\phi < 2m_e$  can be evaded, as it is easily possible for such light mediators to decay into neutrinos before primordial nucleosynthesis. The mixing required for such decays is still sufficiently small that it is essentially unconstrained by neutrino beamline experiments like LSND [83].

We calculate the branching ratios for the case of mass mixing by analytically rescaling the branching ratios



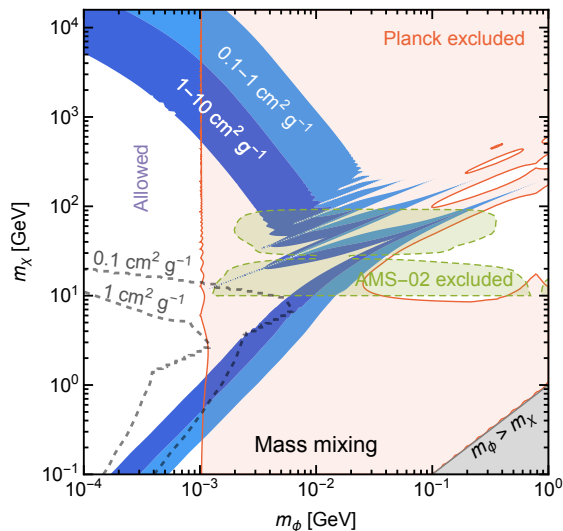


FIG. 4. Constraints on DM particles annihilating into vector mediators that couple to the SM via mass mixing with the  $Z$  boson. The black dashed lines indicate constant values of  $\langle\sigma_T\rangle_{1000}$  corresponding to astrophysically relevant self-interaction cross sections on cluster scales.

for kinetic mixing, see [66]. We find that even for  $m_\chi > 2m_e$  the mediator decays dominantly invisibly up to the point where hadronic resonances become impor-

tant. As a result, all our constraints are significantly weakened (see Fig. 4). Nevertheless, the parameter region required to obtain sizeable self-interaction cross sections is still solidly excluded by CMB constraints, as long as  $m_\phi > 2m_e$ . For smaller mediator masses and  $m_\chi > 1$  TeV it is possible to evade the experimental constraints discussed above and still have self-interactions rates of a phenomenologically relevant strength.

For  $m_\phi \lesssim 1$  MeV and  $m_\chi \lesssim 10$  GeV bounds on the self-interaction cross section at cluster scales become relevant [3, 4, 13–16]. While there is no general consensus regarding the precise value of these bounds (which may even depend on the details of the underlying particle physics), cross sections above  $1 \text{ cm}^2 \text{ g}^{-1}$  are likely excluded, while cross sections below  $0.1 \text{ cm}^2 \text{ g}^{-1}$  are likely consistent with all observations [7–12]. We indicate these two values of  $\langle\sigma_T\rangle_{1000}$  with dashed black lines in Fig. 4 to give an estimate for the strength of these constraints. One can see that cluster constraints are much more easily evaded in the classical regime than in the Born regime, where the velocity dependence of the self-interaction cross section is rather weak. These considerations apply of course in the same way also to Figs. 2 and 3, where the corresponding regions of parameter space are already excluded by other observations and hence the cluster constraints are not shown explicitly.

\* torsten.bringmann@fys.uio.no

† felix.kahlhoefer@desy.de

‡ kai.schmidt-hoberg@desy.de

§ p.s.walia@fys.uio.no

- [1] A. A. de Laix, R. J. Scherrer, and R. K. Schaefer, *Astrophys. J.* **452** (1995), 495, [astro-ph/9502087].
- [2] D. N. Spergel and P. J. Steinhardt, *Phys. Rev. Lett.* **84** (2000), 3760–3763, [astro-ph/9909386].
- [3] M. R. Buckley and P. J. Fox, *Phys. Rev.* **D81** (2010), 083522, [0911.3898].
- [4] A. Loeb and N. Weiner, *Phys. Rev. Lett.* **106** (2011), 171302, [1011.6374].
- [5] J. Zavala, M. Vogelsberger, and M. G. Walker, *Monthly Notices of the Royal Astronomical Society: Letters* **431** (2013), L20–L24, [1211.6426].
- [6] M. Vogelsberger, J. Zavala, and A. Loeb, *Mon. Not. Roy. Astron. Soc.* **423** (2012), 3740, [1201.5892].
- [7] M. Markevitch, A. H. Gonzalez, D. Clowe, A. Vikhlinin, L. David, W. Forman, C. Jones, S. Murray, and W. Tucker, *Astrophys. J.* **606** (2004), 819–824, [astro-ph/0309303].
- [8] S. W. Randall, M. Markevitch, D. Clowe, A. H. Gonzalez, and M. Bradac, *Astrophys. J.* **679** (2008), 1173–1180, [0704.0261].
- [9] A. H. G. Peter, M. Rocha, J. S. Bullock, and M. Kaplinghat, *Mon. Not. Roy. Astron. Soc.* **430** (2013), 105, [1208.3026].
- [10] M. Rocha, A. H. G. Peter, J. S. Bullock, M. Kaplinghat, S. Garrison-Kimmel, J. Onorbe, and L. A. Moustakas, *Mon. Not. Roy. Astron. Soc.* **430** (2013), 81–104, [1208.3025].
- [11] F. Kahlhoefer, K. Schmidt-Hoberg, M. T. Frandsen, and S. Sarkar, *Mon. Not. Roy. Astron. Soc.* **437** (2014), no. 3, 2865–2881, [1308.3419].
- [12] D. Harvey, R. Massey, T. Kitching, A. Taylor, and E. Tittley, *Science* **347** (2015), 1462–1465, [1503.07675].
- [13] L. Ackerman, M. R. Buckley, S. M. Carroll, and M. Kamionkowski, *Phys. Rev.* **D79** (2009), 023519, [0810.5126], [277(2008)].
- [14] J. L. Feng, M. Kaplinghat, H. Tu, and H.-B. Yu, *JCAP* **0907** (2009), 004, [0905.3039].
- [15] L. G. van den Aarsen, T. Bringmann, and C. Pfrommer, *Phys. Rev. Lett.* **109** (2012), 231301, [1205.5809].
- [16] S. Tulin, H.-B. Yu, and K. M. Zurek, *Phys. Rev.* **D87** (2013), no. 11, 115007, [1302.3898].
- [17] P. Gondolo and G. Gelmini, *Nucl. Phys.* **B360** (1991), 145–179.
- [18] X. Chu, C. Garcia-Cely, and T. Hambye, *JHEP* **11** (2016), 048, [1609.00399].
- [19] M. Pospelov, A. Ritz, and M. B. Voloshin, *Phys. Lett.* **B662** (2008), 53–61, [0711.4866].
- [20] M. Kaplinghat, S. Tulin, and H.-B. Yu, *Phys. Rev.* **D89** (2014), no. 3, 035009, [1310.7945].
- [21] E. Del Nobile, M. Kaplinghat, and H.-B. Yu, *JCAP* **1510** (2015), no. 10, 055, [1507.04007].
- [22] C. Kouvaris, I. M. Shoemaker, and K. Tuominen, *Phys. Rev.* **D91** (2015), no. 4, 043519, [1411.3730].
- [23] N. Bernal, X. Chu, C. Garcia-Cely, T. Hambye, and B. Zaldivar, *JCAP* **1603** (2016), no. 03, 018, [1510.08063].

- [24] K. Kainulainen, K. Tuominen, and V. Vaskonen, *Phys. Rev.* **D93** (2016), no. 1, 015016, [1507.04931].
- [25] M. Kamionkowski and S. Profumo, *Phys. Rev. Lett.* **101** (2008), 261301, [0810.3233].
- [26] J. Zavala, M. Vogelsberger, and S. D. M. White, *Phys. Rev.* **D81** (2010), 083502, [0910.5221].
- [27] J. L. Feng, M. Kaplinghat, and H.-B. Yu, *Phys. Rev.* **D82** (2010), 083525, [1005.4678].
- [28] J. Hisano, M. Kawasaki, K. Kohri, T. Moroi, K. Nakayama, and T. Sekiguchi, *Phys. Rev.* **D83** (2011), 123511, [1102.4658].
- [29] J. L. Feng, M. Kaplinghat, and H.-B. Yu, *Phys. Rev. Lett.* **104** (2010), 151301, [0911.0422].
- [30] F.-Y. Cyr-Racine, K. Sigurdson, J. Zavala, T. Bringmann, M. Vogelsberger, and C. Pfrommer, *Phys. Rev.* **D93** (2016), no. 12, 123527, [1512.05344].
- [31] W. J. G. de Blok and S. S. McGaugh, *Mon. Not. Roy. Astron. Soc.* **290** (1997), 533–552, [astro-ph/9704274].
- [32] S.-H. Oh, W. J. G. de Blok, E. Brinks, F. Walter, and R. C. Kennicutt, Jr, *Astron. J.* **141** (2011), 193, [1011.0899].
- [33] M. G. Walker and J. Penarrubia, *Astrophys. J.* **742** (2011), 20, [1108.2404].
- [34] M. Boylan-Kolchin, J. S. Bullock, and M. Kaplinghat, *Mon. Not. Roy. Astron. Soc.* **415** (2011), L40, [1103.0007].
- [35] E. Papastergis, R. Giovanelli, M. P. Haynes, and F. Shankar, *Astron. Astrophys.* **574** (2015), A113, [1407.4665].
- [36] Planck, P. A. R. Ade et al., *Astron. Astrophys.* **594** (2016), A13, [1502.01589].
- [37] A. Sommerfeld, *Annalen der Physik* **403** (1931), 207.
- [38] N. Arkani-Hamed, D. P. Finkbeiner, T. R. Slatyer, and N. Weiner, *Phys. Rev.* **D79** (2009), 015014, [0810.0713].
- [39] S. Cassel, *J. Phys.* **G37** (2010), 105009, [0903.5307].
- [40] R. Iengo, *JHEP* **05** (2009), 024, [0902.0688].
- [41] T. R. Slatyer, *JCAP* **1002** (2010), 028, [0910.5713].
- [42] J. B. Dent, S. Dutta, and R. J. Scherrer, *Phys. Lett.* **B687** (2010), 275–279, [0909.4128].
- [43] L. G. van den Aarssen, T. Bringmann, and Y. C. Goedecke, *Phys. Rev.* **D85** (2012), 123512, [1202.5456].
- [44] B. von Harling and K. Petraki, *JCAP* **1412** (2014), 033, [1407.7874].
- [45] J. A. Adams, S. Sarkar, and D. W. Sciama, *Mon. Not. Roy. Astron. Soc.* **301** (1998), 210–214, [astro-ph/9805108].
- [46] X.-L. Chen and M. Kamionkowski, *Phys. Rev.* **D70** (2004), 043502, [astro-ph/0310473].
- [47] T. R. Slatyer, N. Padmanabhan, and D. P. Finkbeiner, *Phys. Rev.* **D80** (2009), 043526, [0906.1197].
- [48] S. Galli, F. Iocco, G. Bertone, and A. Melchiorri, *Phys. Rev.* **D80** (2009), 023505, [0905.0003].
- [49] J. M. Cline and P. Scott, *JCAP* **1303** (2013), 044, [1301.5908], [Erratum: *JCAP*1305.E01(2013)].
- [50] H. Liu, T. R. Slatyer, and J. Zavala, *Phys. Rev.* **D94** (2016), no. 6, 063507, [1604.02457].
- [51] T. R. Slatyer, *Phys. Rev.* **D93** (2016), no. 2, 023527, [1506.03811].
- [52] T. Bringmann, *New J. Phys.* **11** (2009), 105027, [0903.0189].
- [53] R. A. C. Croft, D. H. Weinberg, N. Katz, and L. Hernquist, *Astrophys. J.* **495** (1998), 44, [astro-ph/9708018].
- [54] R. A. C. Croft, D. H. Weinberg, M. Bolte, S. Burles, L. Hernquist, N. Katz, D. Kirkman, and D. Tytler, *Astrophys. J.* **581** (2002), 20–52, [astro-ph/0012324].
- [55] T. Bringmann, H. T. Ihle, J. Kersten, and P. Walia, *Phys. Rev.* **D94** (2016), no. 10, 103529, [1603.04884].
- [56] Fermi-LAT, M. Ackermann et al., *Phys. Rev. Lett.* **115** (2015), no. 23, 231301, [1503.02641].
- [57] L. Bergstrom, G. Bertone, T. Bringmann, J. Edsjo, and M. Taoso, *Phys. Rev.* **D79** (2009), 081303, [0812.3895].
- [58] G. D. Martinez, *Mon. Not. Roy. Astron. Soc.* **451** (2015), no. 3, 2524–2535, [1309.2641].
- [59] Fermi-LAT, M. Ackermann et al., *Phys. Rev.* **D89** (2014), 042001, [1310.0828].
- [60] AMS Collaboration, L. Accardo et al., *Phys. Rev. Lett.* **113** (2014), 121101.
- [61] AMS Collaboration, M. Aguilar et al., *Phys. Rev. Lett.* **113** (2014), 121102.
- [62] L. Bergstrom, T. Bringmann, I. Cholis, D. Hooper, and C. Weniger, *Phys. Rev. Lett.* **111** (2013), 171101, [1306.3983].
- [63] G. Elor, N. L. Rodd, T. R. Slatyer, and W. Xue, *JCAP* **1606** (2016), no. 06, 024, [1511.08787].
- [64] M. Pospelov and A. Ritz, *Phys. Lett.* **B671** (2009), 391–397, [0810.1502].
- [65] D. Feldman, B. Kors, and P. Nath, *Phys. Rev.* **D75** (2007), 023503, [hep-ph/0610133].
- [66] M. T. Frandsen, F. Kahlhoefer, S. Sarkar, and K. Schmidt-Hoberg, *JHEP* **09** (2011), 128, [1107.2118].
- [67] R. Foot, *Int. J. Mod. Phys.* **D13** (2004), 2161–2192, [astro-ph/0407623].
- [68] J. Redondo and M. Postma, *JCAP* **0902** (2009), 005, [0811.0326].
- [69] D. Curtin, R. Essig, S. Gori, and J. Shelton, *JHEP* **02** (2015), 157, [1412.0018].
- [70] X. Chu, T. Hambye, and M. H. G. Tytgat, *JCAP* **1205** (2012), 034, [1112.0493].
- [71] J. Redondo and G. Raffelt, *JCAP* **1308** (2013), 034, [1305.2920].
- [72] C. Garcia-Cely and J. Heeck, *JCAP* **1608** (2016), 023, [1605.08049].
- [73] IceCube, R. Abbasi et al., *The IceCube Neutrino Observatory IV: Searches for Dark Matter and Exotic Particles*, in *Proceedings, 32nd International Cosmic Ray Conference (ICRC 2011): Beijing, China, August 11-18, 2011*, 2011.
- [74] B. Dasgupta and R. Laha, *Phys. Rev.* **D86** (2012), 093001, [1206.1322].
- [75] IceCube, M. G. Aartsen et al., *Eur. Phys. J.* **C75** (2015), no. 10, 492, [1505.07259].
- [76] C. El Aisati, M. Gustafsson, and T. Hambye, *Phys. Rev.* **D92** (2015), no. 12, 123515, [1506.02657].
- [77] T. Bringmann, J. Hasenkamp, and J. Kersten, *JCAP* **1407** (2014), 042, [1312.4947].
- [78] B. Dasgupta and J. Kopp, *Phys. Rev. Lett.* **112** (2014), no. 3, 031803, [1310.6337].
- [79] X. Chu and B. Dasgupta, *Phys. Rev. Lett.* **113** (2014), no. 16, 161301, [1404.6127].
- [80] P. Ko and Y. Tang, *Phys. Lett.* **B739** (2014), 62–67, [1404.0236].
- [81] T. Binder, L. Covi, A. Kamada, H. Murayama, T. Takahashi, and N. Yoshida, *JCAP* **1611** (2016), no. 11, 043, [1602.07624].
- [82] J. L. Feng, H. Tu, and H.-B. Yu, *JCAP* **0810** (2008), 043, [0808.2318].
- [83] P. deNiverville, M. Pospelov, and A. Ritz, *Phys. Rev.* **D84** (2011), 075020, [1107.4580].